

Aarno Palotie, Mari Kaunisto, Jarmo Harju, Kimmo Pitkänen, Markus Perola ja Anu Jalanko

## FinnGen-tutkimuksen lupaukset

FinnGen-tutkimushankkeen tavoitteena on lisätä ymmärtämystämme tavallisten sairauksien geneettisestä taustasta hyödyntämällä 500 000 suomalaisen genomitietoa ja terveystietoa kerättyä sairaustietoa. Kyseessä on esikilpailullinen, tieteellinen tutkimushanke, jonka osapuolina ovat suomalaiset yliopistosairaanhoitopiirit, yliopistot, THL, Veripalvelu ja yhdeksän kansainvälistä lääkeyritystä. Tutkimuksessa haetaan assosiaatioita geenivarianttien ja sairauksien välillä, tavoitteena muun muassa suomalaisiin rikastuneiden koodaavien geenivarianttien tunnistaminen. Tällaiset variantit voivat auttaa sairausmekanismien selvittämisessä tarjoamalla ohituskaistan toiminnallisille jatkotutkimuksille ja siten biologiselle ymmärrykselle. FinnGenin tuloksia pyritään myös yhdistämään muiden suurten kansainvälisten biopankkihankkeiden kanssa. Näissä meta-analyyseissä FinnGenin tuottamia tuloksia voidaan vahvistaa ja kopioida ja tehdä globaaleja suuranalyyskejä. FinnGen on merkittävästi auttanut biopankkien käynnistysvaihetta ja sen stimuloimat jatkohankkeet osaltaan vahvistavat suomalaista tutkimusinfrastruktuuria. Hanke myös hioo genomitutkimuksen hallinnollisia käytäntöjä tunnistamalla nykyisen säädösympäristön kipukohtia ja voi näin toivottavasti osaltaan auttaa uusien säädösten valmistelutyötä.

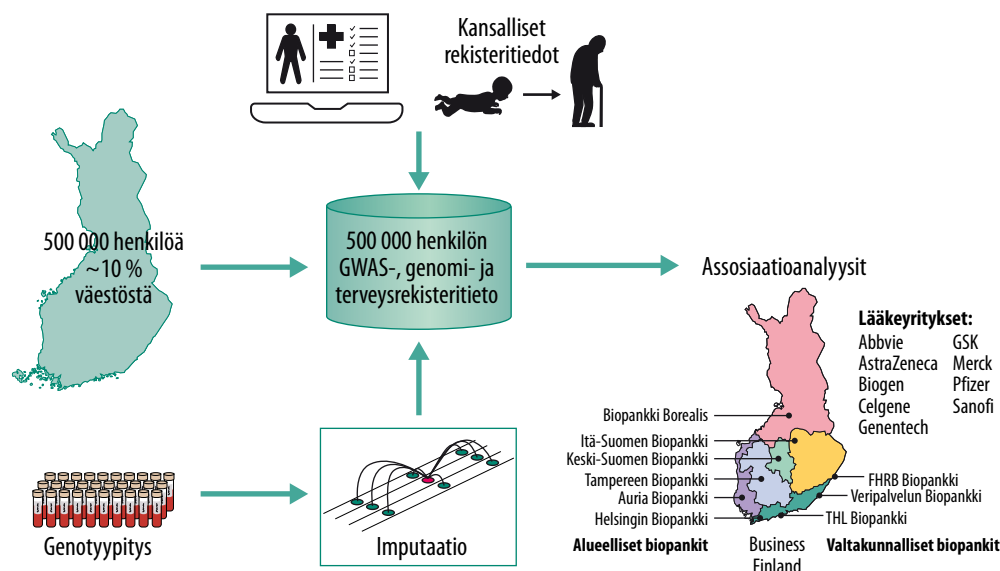
**F**innGen-tutkimushankkeen tavoitteena on lisätä tietoa sairausmekanismeista tutkimalla 500 000 suomalaisen genomi- ja terveystietoja (1). Hankkeen osapuolina ovat suomalaiset biopankit taustaorganisaatioinaan yliopistosairaanhoitopiirit, Helsingin, Turun, Oulun, Tampereen, Jyväskylän ja Itä-Suomen yliopistot, THL ja Veripalvelu yhteistyössä yhdeksän kansainvälisen lääkeyrityksen kanssa. Terveystietojen perustana ovat suomalaiset terveystietorekisterit, ja genomitiedot analysoidaan suomalaisten biopankkien DNA-näytteistä. Hanke käynnistyi elokuussa 2017, ja sen on arvioitu kestävän yhteensä noin kymmenen vuotta. Viidensadan tuhannen osallistujan määrä uskotaan saavutettavan vuonna 2021. FinnGen-tutkimushankkeen kokonaisuus on esitetty **KUVASSA 1**.

### FinnGen on kansainvälisesti ainutlaatuinen hanke

FinnGen on kansainvälisesti ainutlaatuinen tutkimushanke, johon lähes 10 % väestöstä toi-

vottavasti osallistuu. FinnGen voidaan nähdä osana nykyistä tutkimustrendiä, jossa käytetään hyvin suuria tutkimusaineistoja sairauksien geneettisen taustan selvittämiseksi. Lopullisena tähtäimenä on parempi ja yksilöllisempi hoito sekä sairauksien ehkäisy. Tämän trendin lippulaivahanke on UK Biobank, jossa on 500 000 englantilaisen terveys- ja genomitieto kansainvälisen tutkimusyhteisön käytössä (2,3). UK Biobank onkin muovannut koko alan toimintakulttuuria avoimen tieteen ja tiedon periaatteiden mukaiseksi ja biopankin aineistoilla on tuotettu jo yli 600 tieteellistä julkaisua. Esimerkkejä muista kansainvälisistä julkisista suurhankkeista on esitetty **TAULUKOSSA 1**.

**TAULUKOSSA 2** on listattu kaupallisia biopankkihankkeita. Islannin DeCode-hanke on ollut alan uranuurtaja jo 20 vuotta, ja sen tieteellinen tuotanto on vertaansa vailla (4). Toisaalta yksityisenä yrityksenä sen toimintaperiaate on erilainen, eikä se jaa tietovarantojaan tiedeyhteisön käyttöön. Vielä suurempia hankkeita edustavat suoraan kuluttajille myyvät geenitestiyhtiöt kuten 23&Me (5) sekä Ancestry (6),



**KUVA 1.** FinnGen-tutkimushankkeen kokonaisuus.

GWAS = genominlaajuinen assosiaatiotutkimus, FHRB = hematologinen biopankki

joilla molemmilla on jo miljoonien ihmisten geenitiedot yhdistettynä kyselyperusteisiin terveystietoihin. Kuluttajille palveluitaan myyvät yritykset veloittavat ensin näytteenantajia, jalostavat kerätyn tiedon, ja myyvät sitten sen eteenpäin esimerkiksi teollisuudelle (7). Toimintaperiaate on siis varsin erilainen kuin julkisissa hankkeissa.

## FinnGen-tutkimuksen aineisto

FinnGen-tutkimuksen aineisto koostuu toisaalta jo aiemmin kerätystä, lähinnä THL:n Biopankissa olevista väestö- ja sairauskokoelmista, kuten Finrisk/FinTerveys, Terveys 2000, Botnia ja Kaksostutkimusaineisto. Niiden kokonaismäärä tulee olemaan noin 200 000 henkilöä. Toinen kokonaisuus ovat etenevät biopankkeihin kerättävät näytteet, joita kerätään yhteensä noin 300 000. Tässä ovat keskiössä sairaalabiopankit ja Veripalvelun biopankki.

Sairaalabiopankkeihin kertyy ymmärrettävästi näytteitä henkilöistä, jotka käyvät sairaalassa jonkun sairauden takia. Toisaalta aiemmin kerättyjen aineistojen ikäjakauma painottuu vanhoihin ikäluokkiin, joissa sairastavuus on

suurta. Tämä piirre erottaa FinnGen-tutkimuksen useimmista kansainvälisistä suurhankkeista, joissa otos perustuu usein työikäiseen väestöön. Muita FinnGenin erityisvahvuuksia ovat kansallisten rekisteriaineistojen tuottama elämänmittainen terveystieto, suomalaisen väestön geneettinen tausta ja osallistujien sairausdiagnoosien huomattavan suuri määrä.

Biopankkinäytteistä tehdään genominlaajuisen genotyyppitys tarkoitusta varten räätälöidyllä sirulla, jossa on noin 700 000 geenimerkkiä (8). Genotyyppitieto imputoidaan eli täydennetään laskennallisesti suomalaisten referenssi-sekvenssien avulla (9). Käytännössä tämä tarkoittaa, että genotyyppityksen avulla tuotetun ”rungon” ja aiemmin sekvensoitujen suomalaisten genomien avulla voimme hyvin suurella tarkkuudella päätellä, mitä variantteja kukin yksilö kantaa. FinnGenin tällä hetkellä käyttämä referenssitieto perustuu yli 4 000 kokogenomisekvensoituun suomalaiseen. Imputaation jälkeen kustakin näytteestä voidaan luotettavasti analysoida noin 16 miljoonaa geenivarianttia. Myös melko harvinaisten varianttien tautiassoitaatioita voidaan siis tutkia, vaikkei niitä olisi-kaan suoraan genotyyppitetty.

**TAULUKKO 1.** Esimerkkejä julkisesti rahoitetuista kansainvälisistä genomi- ja biopankkitutkimushankkeista.

Hanke	Aineisto	Keräysvaihe	Tietosisältö	Linkki
UK Biobank	500 000 työikäistä englantilaista	Olemassa	Runsaasti terveys- ja mittaustietoa	<a href="https://www.ukbiobank.ac.uk/">https://www.ukbiobank.ac.uk/</a>
All of US	Miljoona yhdysvaltalaisista	Kehitteillä	Runsaasti terveys- ja mittaustietoa	<a href="https://allofus.nih.gov/">https://allofus.nih.gov/</a>
Million Veterans Project	Miljoona yhdysvaltalaisista puolustusvoimien työntekijää	Olemassa, laajennetaan	Runsaasti terveys- ja mittaustietoa	<a href="https://www.research.va.gov/mvp/">https://www.research.va.gov/mvp/</a>
Estonian Biobank	Viron kansallinen biopankki, yli 150 000 osallistujaa	Olemassa, laajennetaan	Runsaasti terveys- ja mittaustietoa, linkki terveystietorekistereihin	<a href="https://www.geenivaramu.ee/en">https://www.geenivaramu.ee/en</a>
Biobank Japan	Yli 200 000 osallistujaa useasta sairaalasta	Olemassa	Sairaus- ja mittaustietoa	<a href="https://biobankjp.org/cohort_3rd/english/">https://biobankjp.org/cohort_3rd/english/</a>
Danish National biobank	Muun muassa 2,2 miljoonan vastasyntyneen seurantaverinäytteen keräys	Olemassa	Linkki terveystietorekisteritietoon	<a href="https://biopeople.eu/fileadmin/user_upload/Editor/Files/Prec-Med2018/Konradsen_presentation.pdf">https://biopeople.eu/fileadmin/user_upload/Editor/Files/Prec-Med2018/Konradsen_presentation.pdf</a>
BioVU, Vanderbilt DNA Databank	Yli 250 000 sairaalapotilasta	Olemassa, kasvaa	Sairaalatiedostot	<a href="https://victor.vanderbilt.edu/pub/biovu/">https://victor.vanderbilt.edu/pub/biovu/</a>

FinnGen-tutkimuksen fenotyyppitieto perustuu pääosin kansallisista terveystietorekistereistä saatuihin tietoihin. Keskeisimmät tietolähteet ovat THL:n sairaala- ja erityispoliklinikkapoitorekisterit ja syöpärekisteri, Kelan lääkeosto- ja korvausrekisterit, Väestörekisteri ja Tilastokeskus. Rekisteritiedot kootaan, yhdistetään ja muokataan THL:ssä niin, että voidaan luoda mielekkäitä sairauspäätöspisteitä, ja samalla yksilön tunnistettavuus vähenee. Rekisteritietojen erityisvahvuuksia ovat niiden kansallinen kattavuus sekä sairaanhoitojärjestelmissä että koko elämänsä ajan. Rajoituksena on, että niistä emme voi luotettavasti tunnistaa keskeisimpiä riskitekijöitä, kuten tupakointia ja ylipainoa. Toinen rajoite on, että joidenkin diagnoosien luotettavuus vaihtelee. Jotkut erikoissalat kirjaavat diagnoosit huolellisesti, joidenkin kohdalla kirjaamiseen ei suhtauduta yhtä suurella huolella. Yliopistosairaaloiden senioriklinikoista koostuvat FinnGenin kliiniset työryhmät ovat keskeisiä rekisteritietojen optimaalisen käytön suunnittelussa.

## Tieteelliset analyysit

FinnGeniin kertyvää genotyyppi- ja fenotyyppitietoa käytetään geenivarianttien sairaus-

yhteyksien tunnistamiseen, niiden vaikutuksen arviointiin ja sairauksien välisten yhteisten mekanismien selvittämiseen. Perustyökaluina ovat genomilaajuiset assosiaatioanalyysit, PheWAS (yksittäisten varianttien fenotyyppiassosiaatioanalyysit) ja polygeenisten riskien (PRS, polygenic risk scores) assosiaatioanalyysit (10). Projektissa valmistuu uusi datakokonaisuus joka kuudes kuukausi, siten että osallistujamäärä lisääntyy aina 40–50 000 henkilöllä. Helmikuussa 2019 valmistuneessa kokonaisuudessa on yli 140 000 suomalaisen laatuvarmistetut genomi- ja terveystiedot. Näistä valmistuneet genomilaajuiset assosiaatioanalyysit ja PheWAS-analyysitulokset ovat projektin tutkijoiden nähtävillä tätä tarkoitusta varten kehitetyllä PheWEB-selaimella. Tutkijoiden saatavilla on myös UK Biobankin tulokset ja näiden kahden hankkeen tuloksiin perustuvat meta-analyysit.

## Biopankit FinnGen partnereina

Suomalaiset biopankit ja niiden menestykselliset toiminta ovat yksi FinnGen-hankkeen kivijaloista. Kun FinnGen-tutkimushankkeen valmistelu aloitettiin, pohjana olivat lähinnä THL:n vuosikymmenten aikana kootut näyte- ja tietomateriaalit. Sairaalabiopankkien toiminta oli juuri

TAULUKKO 2. Esimerkkejä kaupallisista biopankkihankkeista.

Hanke	Tarjottu palvelu	Keräysvaihe	Aineisto	Toiminnan tavoite	Linkki
Ancestry	Sukulaisuussuhteiden ja syntyperän selvityspalveluja	Olemassa, kasvaa	Tiedot yli 5 miljoonasta ihmisestä	Geenitestejä suoraan kuluttajalle, taustalla laaja genomitieto, mutta rajoitettu fenotyyppitieto	<a href="https://www.ancestry.com/">https://www.ancestry.com/</a>
deCode Genetics	Sairauksien geneettisen taustan selvittäminen	Olemassa	Tiedot koko Islannin väestöstä, noin 300 000 henkeä	Geneettisen tiedon käyttö sairausmekanismien selvittämiseksi ja lääkekehityksen avuksi	<a href="https://www.decode.com/publications/">https://www.decode.com/publications/</a>
23andMe	Kaupallinen geenitestipalvelu	Olemassa, kasvaa	Tiedot yli 5 miljoonasta ihmisestä	Geenitestejä suoraan kuluttajalle, tietojen käyttö tutkimuksessa ja tiedon myynti yrityksille	<a href="https://research.23andme.com/">https://research.23andme.com/</a>

käynnistymässä tai alkamassa vakiintua. Valmisteluvaiheessa mukana olleet FinnGen-partnerit näkivätkin juuri biopankkien vakiintumattomuuden suurimpana riskinä. Nyt kun hanke on ollut käynnissä yli puolitoista vuotta, ovat kaikki biopankit saaneet toimintansa vakiinnutettua ja epäilyt ovat hälventyneet.

FinnGen-tutkimus on keskeinen biopankkien tietovarantojen kannalta, sillä kaikki projektissa tuotettu genomitieto palautuu biopankkien omistukseen. Kukin yksittäinen biopankki on liian pieni ollakseen kansainvälisesti merkittävä, mutta kokonaisuutena Suomi on varteenotettava kansainvälinen kumppani. Biopankit ovat FinnGenin ensimmäisen 17 kuukauden aikana keränneet yli 120 000 näytettä, ja keräysvauhti on vakiintunut noin 6 000 näytteeseen kuukaudessa, mikä on selvästi nopeampaa kuin alkuperäiset tavoitteemme. Tämä on vakuuttanut myös teolliset partnerit. Toki 6 000 henkilöä kuukaudessa on varsin pieni osa erikoissairaanhoidon palveluiden käyttäjistä, eli potentiaalia on vielä runsaasti.

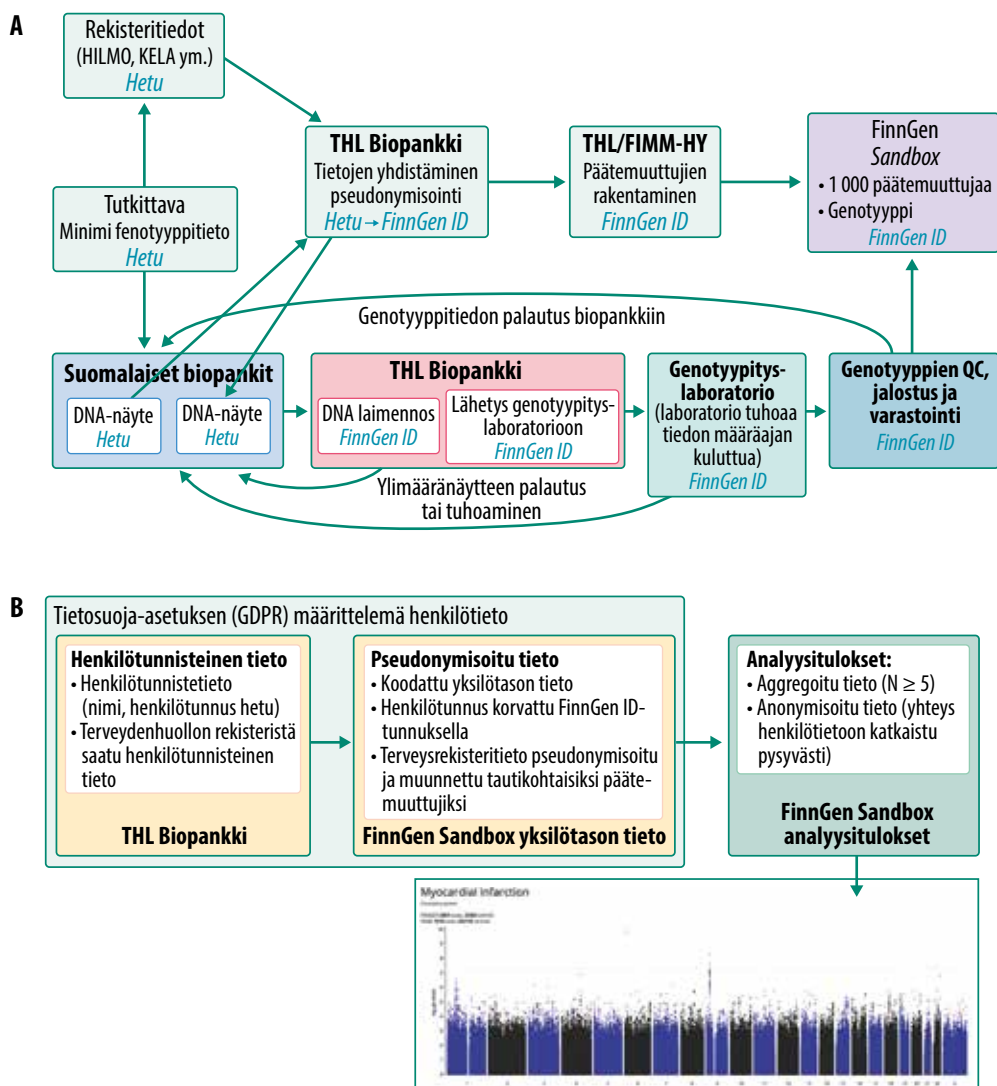
Haasteena on ollut perustietojen ammentaminen sairaalajärjestelmistä. Analyysijä varten keskeisimpien riskitekijöiden, kuten painoindexin ja tupakointitietojen, saaminen ei ole helppoa. Tämä kuvastaa, kuinka hajanaisista tietä sairaalajärjestelmissä on. Lähivuosien aikana tällä osa-alueella on kuitenkin odotettavissa merkittävää kehitystä. Useat sairaalat ovat rakentaneet tietoallasprojekteja ja muita

strategioita, joiden avulla kertynyttä tietoa voisi käyttää paremmin sekä potilashoidossa että tutkimuksessa.

## Teollisuus FinnGen partnereina

Kansainvälinen lääketeollisuus on keskeinen osa FinnGen-hanketta. Teollisuuden osuus rahoituksesta on yli 70 %, ja loppuosa budjetista tulee Business Finlandin kautta. Vaikka lääkeryritysten osuus rahoittajana on keskeinen, kyseessä on tutkimusyhteistyö, jonka tavoitteena on yleisen tiedon lisääminen koko tutkimuskentässä, ei vain FinnGen-konsortion sisällä. Teollisten partnereiden tieteellisesti ansioituneet tutkijat ovat aktiivisesti mukana yhteisönä, yli yritysrajojen, suunnittelemassa ja toteuttamassa hanketta.

Tämä ainutlaatuinen yhteistyömalli on mahdollista siksi, että FinnGen on puhtaasti tutkimukseen keskittyvä hanke. Tähtäimenä on yleisen sairausbiologian ymmärryksen lisääminen tavalla, joka olisi yksittäisille lääkeyrityksille liian kallista ja monessa mielessä mahdotonta. Tavoitteena on, että lisäämällä geneettistä ymmärrystä sairauksien syistä lääkkeiden kehityspolkua voidaan nopeuttaa. On myös näyttöä siitä, että geneettisen tiedon hyödyntäminen lääkekohteen valinnassa vähentää jatkovaiheiden epäonnistumisia (11). FinnGen-tutkimuksessa ei odoteta syntyvän IP-oikeuksien kannalta mielenkiintoista suojattavaa. Jos kuitenkin



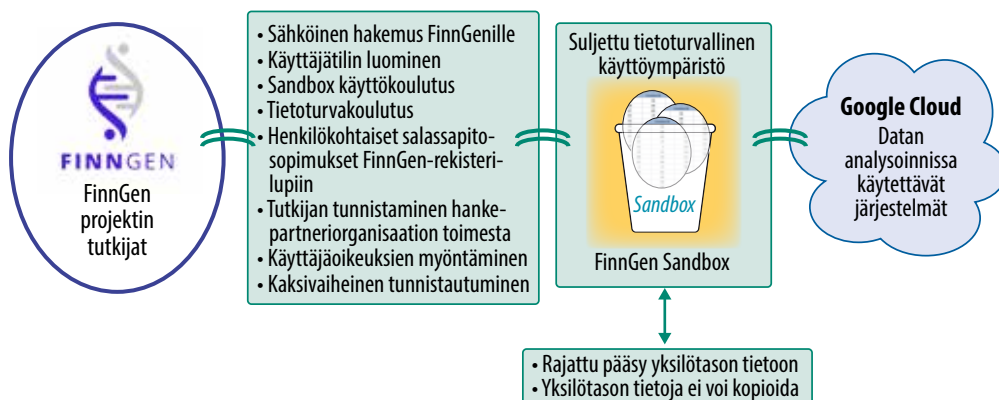
**KUVA 2.** FinnGen-tutkimushankkeen näytteiden ja tietojen kulku (A) ja henkilötietojen käsittelyn yksityiskohdat (B). FinnGenin analyysiryhmä suorittaa yhdessä sovitut analyysit ja valmistelee niiden tulokset hankkeen tutkijoitten käyttöön. QC = quality control, ID = tunnistus, Hetu = henkilötunnus

sellaista ilmenee, olisi omistusoikeus kaikkien partnereiden yhteinen.

## FinnGen näytteiden ja tietojen kulku ja tietoturva

FinnGen-tutkimushankkeessa käytetään suomalaisiin biopankkeihin pääasiassa biopankki-suostumuksen perusteella kerättyjä, verinäytteestä eristettyjä DNA-näytteitä ja niihin liitettäviä tietoja. THL pseudonymisoi kaikkien

biopankkien näytteet, ja ne lähetetään FinnGen ID-tunnuksilla genotyyppityslaboratorioon (ThermoFisher Yhdysvalloissa). Genotyyptystiedot palautuvat Helsingin yliopistoon ja genotyyppityslaboratorioon ei jätetä tiedoista pysyvää kopiota. Näytteiden ja tietojen kulku on esitetty **KUVASSA 2**. THL muokkaa rekisteritiedot tutkimushankkeessa käytettävään muotoon, ja muokatut päätemuuttujatiedot yhdistetään genomitietoon FinnGen ID-tunnisteen perusteella. Tästä syntyy FinnGenin tietova-



**KUVA 3.** FinnGen-tutkimushankkeen analyysiympäristö.

ranto, jonka perusteella tutkimusanalyysit tehdään. Tietovarannossa on ainoastaan pseudonymisoitua tietoa, ja henkilötietoihin pääsyä hallitaan voimassaolevan lainsäädännön vaatimusten mukaisesti (muun muassa tietosuojasetus GDPR) (**KUVA 2**). Nimetyt FinnGen-tutkijat saavat monivaiheisen hakuprosessin jälkeen pääsyn FinnGenin suljettuun käyttöympäristöön (**KUVA 3**). FinnGen-analyyysiryhmä valmistelelee genomilaajuisen assosiaatio-tutkimuksen ja PheWAS:n analyysitulokset FinnGen-tutkijoiden käyttöön.

Suljettu tietoturvallinen käyttöympäristö, FinnGen Sandbox, sijaitsee Googlen pilvipalvelussa (12) (**KUVA 3**). FinnGenin tutkijat voivat tehdä kevyitä analyyseja virtuaalikoneella ja tarkastella tietoa, mutta yksilötason tiedon siirtäminen suljetun käyttöympäristön ulkopuolelle on estetty. Suljetun tietoturvallisen käyttöympäristön rakentaminen ja auditointi on ollut erittäin merkittävä saavutus, ja tekninen ratkaisu on jo nyt saavuttanut suurta kiinnostusta kansainvälisesti. Halutessaan tutkijat voivat siirtää anonymisoituja analyysituloksia ympäristön ulkopuolelle.

## Tutkittavien tiedottaminen tuloksista

Biopankkilain mukaan tutkittavalla on oikeus kysyä biopankista omaan näytteeseensä liittyvistä tutkimuksista. FinnGenin kaltaisessa suu-

ressa geenitutkimushankkeessa luodaan yhteistyössä FINBB:n ja biopankkien kanssa yhtenäiset käytännöt mahdollisten lääketieteellisesti merkittävien sattumalöydösten palauttamisesta tutkittavalle. Käytäntöjen laatimisessa hyödynnetään muun muassa tietoa kansainvälisistä ja kansallisista tutkimuksista, joissa tutkittaville on suoraan palautettu terveydellisesti merkittävää tutkimustietoa. Lisäksi hyödynnetään erillisiä määrittely- ja tutkimushankkeita, joissa selvitetään, minkälaisen geenivarianttien tiedosta on osallistujille todellista ennusteellista terveydellistä hyötyä, sekä suunnitellaan ja testataan sattumalöydösten palauttamista tutkittaville. Asiaan liittyvä pilottitutkimus on kuvattu tässä samassa numerossa (Perola ym. P5-lääketiede jalkautuu Suomeen), ja suunnitteilla oleva kuvattu pilotin jatkotutkimus tulee tutkimaan myös kliinisesti merkittävien varianttien palautuksen kansantaloudellista kokonaisvaikutusta. On myös olemassa kansainvälisiä suosituksia niistä geeneistä ja mutaatioista, joiden kanta-juutta geenitutkimusaineistoissa tulisi selvittää ja mistä tulisi tutkituille antaa palautetta (13,14). Näiden geenivarianttien terveysvaikutuksia ei kuitenkaan ole tutkittu ainakaan suomalaisessa väestössä ja tieto geenivariantteista, joiden vaikutus sairastumisennusteeseen on epäselvä, voisi hämmäntää tutkittavia (15). On eettisesti perusteltua aktiivisesti palauttaa tutkittaville vain sellaista tietoa, jolla on merkitystä.

## Mitä FinnGenin tuloksena voidaan odottaa?

Suomalainen väestö polveutuu pienestä perustajajoukosta ja on kasvanut viimeisen vajaan kolmensadan vuoden aikana nopeasti, useita geneettisiä pullonkauloja läpikäyden. Siksi väestöömme on rikastunut geenivariantteja, jotka ovat muissa väestöissä huomattavasti harvinaisempia (9,16). Rikastuneissa varianteissa on erityisesti koodaavia, geenituotteen toimintaan vahvasti vaikuttavia "loss of function-" ja aminohapon muuttavia muutoksia (9). Osa näistä varianteista assosioituu sairauksiin, joko sairausriskiä lisäten tai vähentäen. Yksi keskeinen motivaatio FinnGen-hankkeelle on näiden suomalaisiin rikastuneiden geenivarianttien tunnistaminen, sillä niiden toivotaan helpottavan geeniassoosiaatioiden biologista tulkintaa.

Yksi genomilajajuisten assosiaatiotutkimusanalyysien haaste on, että assosiaatiosignaalit osuvat usein genomien koodaamattomille alueille, jolloin geenilokuksen assosiaatiosta on pitkä matka toiminnan ymmärtämiseen. Koodaavat variantit saattavat tarjota oikopolun biologisen mekanismin ymmärtämiseksi. Ensimmäiset, alustavat FinnGen-analyysit ovat jo tuoneet vahvistusta tälle hypoteesille. Analyysissä on keskitytty niihin sairauksiin, joissa tämänhetkisen 146 000 osallistujan aineistossa oli riittävästi tapauksia. Kiinnostavia Suomeen rikastuneita, muualla harvinaisia variantteja havaittiin glaukoomassa, tyypin 2 diabeteksessa, tulehduksellisissa paksusuolisairauksissa, astmassa ja uniapneassa. Löydökset vaativat vielä vahvistuksia, mutta antavat vahvaa näyttöä siitä, että lähestymistapamme on mielekäs. Osallistujamäärien kasvaessa voimme odottaa lisää löydöksiä, joiden avulla voimme päästä uusien toiminnallisten mekanismien jäljille.

## Yhteistyö samankaltaisten tutkimushankkeiden kanssa

FinnGen-hanke on maailmanlaajuisestikin merkittävä ja ainutlaatuinen hanke, mutta se ei yksin riitä. Genomivariaatioiden sairausyhteyksien tutkiminen vaatii todella suuria aineistoja. Useiden tutkimuskysymysten kohdalla mikään

yksittäinen, suurikaan aineisto ei yksin pysty vastaamaan tähän tarpeeseen. Siksi suurten hankkeiden tietojen ja tulosten yhdistäminen on välttämätön seuraava askel. Tieteellisesti tehokkainta olisi, jos alkuperäistä, yksilötason tietoa voisi yhdistää. Tämä on kuitenkin harvoin mahdollista henkilötietosuojan ja maiden välisten säädösympäristöjen erilaisuuksien vuoksi. Eri aineistojen meta-analyysit, jossa yksilötason tietoa ei tarvitse jakaa, ovat helpommin toteutettavissa. Jo nyt FinnGen yhdistää automaattisesti vastaavien analyysien tulostiedot UK Biobankin analyyyseistä. Tämä tarjoaa mahdollisuuden tulosten replikoimiseen, yksittäisissä aineistoissa liian heikkojen signaalien vahvistamiseen sekä signaalialueen täsmentämiseen väestöjen erilaisen kytkentäepätasapainorakenteen avulla. Alustavat löydökset ovat lupaavia, ja tavoitteenamme onkin rakentaa toimiva yhteistyö useiden merkittävien genomitutkimushankkeiden kanssa.

## Säädösympäristö

Keskeinen edellytys FinnGen-hankkeen käynnistymiselle ja toiminnalle on ollut mahdollista suomalainen säädösympäristö, erityisesti Biopankkilaki. FinnGen-hankkeen valmistelun ja alkuvaiheen aikana on odoteltu uutta selkiyttävää lainsäädäntöä, jossa rekisteritiedon toisiokäytön ja genomikeskuksen toiminnat määritettäisiin sekä nykyisen biopankkilain haasteita korjattaisiin. Hiljattain eduskunnassa hyväksytty rekisteritiedon toisiolaki toivottavasti osaltaan selkeyttää nykyisiä käytäntöjä. Lisäksi FinnGen-hankkeen aikana astui voimaan uusi eurooppalainen tietosuoja-asetus (GDPR). Tämä muutosten odottelu ja toimintaympäristön epävarmuus ovat värittäneet FinnGenin käytännön toimintaa.

Kuten yllä on kuvattu, tietosuoja- ja tietoturva-asiat on alusta alkaen otettu erityisen vakavasti, ennakoiden GDPR:n voimaantuloa. Tämänhetkinen säädösympäristö on kuitenkin monimutkainen ja yhteensopimaton. Tästä syystä eri viranomaiset tulkitsevat säädöksiä eri tavoin, joskus varsin rajatusti vain yhdestä näkökulmasta. Selkiyttävän ja mahdollistavan lainsäädännön aikaansaaminen olisi tervetul-



## Ydinasiat

- » FinnGen on kansainvälisesti ainutlaatuisen genomi- ja terveystietoa yhdistävä tutkimushanke, johon toivotaan biopankkien kautta 500 000:ta osallistujaa.
- » Tutkimuksen pää tavoitteena on geenivarianttien sairausyhteyksien tunnistaminen muun muassa suomalaisiin rikastuneita koodaavia geenivariantteja hyödyntäen.
- » FinnGen-hanke on edistänyt merkittävästi suomalaisten biopankkien yhteistyötä ja nopeuttanut näytekeraäyksiä.
- » FinnGen-tutkimushankkeessa toteutuu akateemisten ja yritysten tutkijoiden ainutlaatuinen yhteistyö.

lutta tutkijoiden oikeusturvan kannalta. Tutkimuslaitosten ja sairaaloiden on ollut välttämätöntä osoittaa yhä enemmän resursseja tukemaan tutkijoita lainsäädännön tulkinnoissa ja tutkimuksessa tarvittavan lupaprosessien läpiviemisessä. Asiaa ei helpota, että tilanne monimutkaistuu usein lisää kansainvälisissä yhteistyöprojekteissa (17). Vielä tärkeämpää kuitenkin olisi, että näytteensä ja tietonsa tutkimuksen hyväksi luovuttaneiden kansalaisten toive toteutuu eikä säädösten tulkinnan sekavuus estä sen toteutumista.

## Mitä seuraavaksi?

FinnGen-tutkimushankkeen omien tieteellisten tavoitteiden lisäksi FinnGen pyrkii stimuloimaan uusia tutkimus- ja kehitystarpeita sekä biopankeille että kotimaiselle orastavalle teollisuudelle. Osa uusista hankkeista on puhtaasti akateemisia, osa yhteistyötä teollisuuden kanssa. Me kutsumme näitä FinnGenistä stimuloituneita erillisiä hankkeita ”add-on-hankkeiksi” ja niiden organisoimisesta vastaa Biopankkien osuuskunta, FINBB. Toivomme siis, että FinnGenin tuottamat mielenkiintoiset tutkimustulokset stimuloivat jatkohankkeita, joissa esimerkiksi tarvitaan tarkempaa kliinistä,

kliinispatologista tai kuvantamistietoa, tai halutaan kutsua tutkittavia kohdennetusti jatkotutkimuksiin.

FinnGen-hankkeen jatkokehittäminen on mahdollista suunnata tukemaan biopankkien kehittymistä pelkistä näytteentallentajista sairaaloiden näytteiden, kliinisen tiedon ja tutkimustiedon tietopankeiksi. Tämä on mielekäs, kansainvälisesti looginen seuraava kehitysvaihe, jonka ensimmäisiä askeleita edistyneimmät biopankit yhteistyössä sairaaloiden tietoaltaiden kanssa kykenevät jo ottamaan. Kehitysuunta tarvitsee kuitenkin runsaasti lisää osaamista ja panostuksia. Tässäkin on keskeistä muistaa, että kukin yksittäinen suomalainen toimija on kansainvälisessä vertailussa pieni ja syrjäinen. Jos visiota kehitetään yhdessä, pitäen Suomea kokonaisuutena, saatamme hyvinkin pärjätä ja saada terveydenhuollollemme ja sen tutkimukselle uusia mahdollisuuksia.

## Lopuksi

FinnGenin luomat uudet mahdollisuudet tutkimukselle ja sitä kautta terveydenhoidolle ovat varsin kiistattomat. Uusiin teknologioihin ja uusiin mahdollisuuksiin liittyy kuitenkin aina keskustelua, jossa ääripää korostuvat. Toisaalta kuulemme ylödötuksia ja hypetystä: luodaan uutta Nokiaa ja uudistetaan lääketiede kaikilta osin, kohta kaikki selailevat omia genomejaan ja biopankit ovat sairaaloiden rahasampoja. Toisessa ääripäässä ajatellaan, että henkilötiedot ryöstäytyvät kaiken kansan nähtäväksi, yksityisyys on mennyt, ja jokainen naapuri tuntee toistensa genomivariaatiot. Tasapainoisen, syvällisen asiantuntevan ja asiapohjaisen keskustelun saavuttaminen on vaikeaa.

Biopankkien kehittäminen seuraavaan vaiheeseen on sekin vaikeaa. Nyt näytekeraäys ja monet muut perusasiat on saatu kuntoon. Samalla taustaorganisaatiot kuitenkin odottavat, että jo ennestään vaatimatonta perusrahoitusta voitaisiin pienentää ja toimintamenoja kattaa myyntituotoilla. Tämä tuskin on lyhyellä tähtämällä realistista tai viisasta eikä ole onnistunut muuallakaan maailmassa. Ilman biopankkitoiminnan kansallisten tavoitteiden kirkastamista ja tavoitteiden mukaisia määrätietoista inves-



toiteja ei synny kunnollista lopputulosta, vaan pieniä kituvia yksiköitä. Hyvä esimerkki tulee Tanskasta, jossa kansallisen biopankin (18) pitkäjänteisen rakentamistyön lisäksi valtio on nyt investoinut Tanskan genomikeskukseen 13 miljoonaa euroa (19). Tämä on välittömästi poikunut Novo Nordisk -tutkimussäätiön 133 miljoonan dollarin lisäinvestoinnin (20). Myös Viron hallitus on useana vuonna investoinut

biopankkitoimintaan ja on sitoutunut tähän jatkossakin. Isossa-Britanniassa on käynnistynyt hanke, joka tähtää viiden miljoonan brittiläisen potilaan kokogenomin sekvensoimiseksi (21). Kun maailmalla investoidaan suuresti, on lyhytnäköistä asettaa epärealistisia odotuksia biopankkien tuottavuudesta Suomessa, ilman johdonmukaisia investointeja ja Suomea kokonaisuutena ajatellen. ■

**AARNO PALOTIE, LKT, professori, tutkimusjohtaja**

Suomen molekyyli lääketieteen instituutti (FIMM), HiLIFE, Helsingin yliopisto, Center for Genome Research, Massachusetts General Hospital, The Broad Institute of MIT and Harvard

**MARI KAUNISTO, FT, vanhempi tutkija**

**JARMO HARJU, DI, tietojärjestelmäprojektipäällikkö**

Suomen molekyyli lääketieteen instituutti (FIMM), HiLIFE, Helsingin yliopisto

**ANU JALANKO, TkT, dosentti, vanhempi tutkija**

Suomen molekyyli lääketieteen instituutti (FIMM), HiLIFE, Helsingin yliopisto, Genomi- ja biomarkkeri yksikkö, THL

**KIMMO PITKÄNEN, FT, laboratoriopäällikkö**

Helsingin biopankki, Hus

**MARKUS PEROLA, LT, tutkimusprofessori, yksikön päällikkö**

Genomi- ja biomarkkeri yksikkö, THL

**VASTUUTOIMITTAJA**

Tuomas Mirtti

**SIDONNAISUUDET**

**Aarno Palotie:** Osa palkasta FinnGen-hankkeesta, jota rahoittavat Business Finland ja yhdeksän lääkeyritystä (Abbvie, AstraZeneca, Biogen, Celgene, Genentech, GSK, Merck/MSD, Pfizer ja Sanofi)

**Mari Kaunisto:** Osa palkasta FinnGen-hankkeesta, jota rahoittavat Business Finland ja yhdeksän lääkeyritystä (Abbvie, AstraZeneca, Biogen, Celgene, Genentech, GSK, Merck/MSD, Pfizer ja Sanofi), luento-/asiantuntijapalkkio (Lääketietokeskus)

**Jarmo Harju:** Ei sidonnaisuuksia

**Kimmo Pitkänen:** Luento-/asiantuntijapalkkio (LabQuality-päivät, Lääketietokeskus, Tekes)

**Perola Markus:** Apuraha (NovoNordisk Foundation apuraha työryhmälle 2015), osallistuminen tutkijana Business Finlandin ja yhdeksän lääkeyrityksen (Abbvie, AstraZeneca, Biogen, Celgene, Genentech, Merck/MSD, Pfizer, GSK ja Sanofi) rahoittamaan FinnGen-hankkeeseen

**Anu Jalanko:** Osa palkasta FinnGen-hankkeesta, jota rahoittavat Business Finland ja yhdeksän lääkeyritystä (Abbvie, AstraZeneca, Biogen, Celgene, Genentech, GSK, Merck/MSD, Pfizer ja Sanofi)

**SUMMARY**

**The promises of FinnGen research**

The FinnGen research project will construct a unique resource of 500 000 Finns that enables ambitious study designs to improve our understanding of the genetic background of diseases. The FinnGen project is an academic-pharma partnership that involves nine Finnish biobanks, all Finnish University Hospitals and their respective Universities, the National Institute for Health and Welfare (THL), the Finnish Red Cross Blood Service and nine large pharmaceutical companies. The study aims to produce close to complete genome variant data from all the 500 000 participants combined with the extensive longitudinal health register data available on all Finns. This data provides unique opportunities to study disease associations (GWAS and PheWAS), disease trajectories and comorbidities.

## KIRJALLISUUTTA

1. FinnGen-tutkimushanke [verkkosivu]. [www.finnngen.fi](http://www.finnngen.fi).
2. Sudlow C, Gallacher J, Allen N, ym. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;12:e1001779.
3. Bycroft C, Freeman C, Petkova D, ym. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562: 203–9.
4. deCODE genetics [verkkosivu]. <https://www.decode.com/publications/>.
5. 23andMe [verkkosivu]. <https://research.23andme.com/>.
6. AncestryDNA Research and Collaboration [verkkosivu]. [www.ancestry.com/cs/collaborations](http://www.ancestry.com/cs/collaborations).
7. Petrone J. 23andMe wades further into drug discovery. *Nat Biotechnol* 2017;35: 897.
8. Genotyping [verkkosivu]. FinnGen. [www.finnngen.fi/en/researchers/genotyping](http://www.finnngen.fi/en/researchers/genotyping).
9. Lim ET, Würtz P, Havulinna AS, ym. Sequencing Initiative Suomi (SiSu) Project. Distribution and medical impact of loss-of-function variants in the Finnish founder population. *PLoS Genet* 2014; 10:e1004494.
10. Bush W, Oetjens MR, Crawford DC. Unravelling the human genome – phenome relationship using phenome-wide association studies. *Nat Rev Genet* 2016;17: 129–45.
11. Scott RA, Freitag DF, Li L, ym. A genomic approach to therapeutic target validation identifies a glucose-lowering GLP1R variant protective for coronary heart disease. *Sci Transl Med* 2016;8:341ra76.
12. Google Cloud [verkkosivu]. <https://cloud.google.com/>.
13. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing [verkkotietokanta]. ClinVar. [www.ncbi.nlm.nih.gov/clinvar/docs/acmg/](http://www.ncbi.nlm.nih.gov/clinvar/docs/acmg/).
14. Kalia SS, Adelman K, Bale SJ, ym. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med* 2017;19:249–55.
15. Wong CA, Hernandez AF, Califf RM. Return of research results to study participants uncharted and untested. *JAMA* 2018;320: 435–36.
16. Martin AR, Karczewski KJ, Kerminen S, ym. Haplotype sharing provides insights into fine-scale population history and disease in Finland. *Am J Hum Genet* 2018; 102:760–75.
17. Simell BA, Törnwall OM, Hämäläinen I, ym. Transnational access to large prospective cohorts in Europe: Current trends and unmet needs. *Nat Biotechnol* 2019;49: 98–103.
18. Danish National Biobank [verkkosivu]. [www.danishnationalbiobank.com/](http://www.danishnationalbiobank.com/).
19. The GenomeDenmark platform [verkkosivu]. [www.genomedenmark.dk/english/](http://www.genomedenmark.dk/english/).
20. Novo Nordisk Foundation supports new national genome centre [verkkosivu]. <https://novonordiskfonden.dk/en/news/novo-nordisk-foundation-supports-new-national-genome-centre/>.
21. Matt Hancock announces ambition to map 5 million genomes. GOV.UK 2.10.2018. <https://www.gov.uk/government/news/matt-hancock-announces-ambition-to-map-5-million-genomes>.